Comparison of methods of algebraic strain estimation from Rf/ϕ data: A unified theory of 2D strain analysis

Yamaji, Atsushi

Comparison of methods of algebraic strain estimation from Rf/ϕ data: A unified theory of 2D strain analysis

Atsushi Yamaji[a,*]

[a]*Division of Earth and Planetary Sciences, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan*

Abstract

A unified development of the subject of the algebraic strain analysis methods using Rf/ϕ data is outlined, embodying the main features the theories of Shimamoto and Ikeda, Mulchrone et al. and Yamaji. It is shown that the theories yields an identical strain ellipse from the same data set. However, error estimation in that of Shimamoto and Ikeda is difficult owing to the distortion of its parameter space: Resolution of their method depends on the choice of a reference orientation in the plane where strain markers are observed. In this respect, the remaining two theories have advantages. The hyperbolic vector mean method was developed in the Minkowski 3-space, thereby linked seamlessly with the visualizing methods of Rf/ϕ data, optimal strain and its confidence region. In addition, the residuals of the optimal strain ellipse determined by this method have clear physical meanings concerning logarithmic strains needed to transform a unit circle to given ellipses.

Keywords: strain analysis, hyperbolic geometry, Minkowski space, Lorentz transformation

1. Introduction

Assuming homogeneous strain for all strain markers and their matrix, Rf/ϕ strain analysis determines the aspect ratio and major-axis orientation of strain ellipse from the markers. Isotropic pre-strain shape fabric is usually assumed for this purpose (e.g., Ramsay, 1967; Ramsay and Huber, 1983; Lisle, 1985; Mulchrone et al., 2003). Matthews et al. (1974) pioneered in developing algebraic methods for the analysis. That is, they found that the component-wise mean of the matrices representing post-strain ellipses gave the aspect ratio of strain ellipse, though it required an independent estimate of the maximum stretching orientation. This limitation was removed by Shimamoto and Ikeda (1976). A few methods were proposed to deal with Rf/ϕ with not only isotropic pre-strain fabric but also the special types of anisotropic one (Dunnet and Siddans, 1971; Yamaji, 2005, 2008).

Three algebraic methods have been proposed for Rf/ϕ strain analysis (Shimamoto and Ikeda, 1976; Mulchrone et al., 2003; Yamaji, 2008). The computer programs for the algebraic methods are available at the websites of K. F. Mulchrone, F. W. Vollmer and A. Yamaji. They yield accurate solutions rapidly by taking the simple means of the quantities that represent elliptical strain markers. Yamaji (2008) showed that his hyperbolic vector mean method and the method of Mulchrone et al. (2003) results in the same strain ellipse. In this paper, the result of the method of Shimamoto and Ikeda (1976) is shown to be identical as well.

However, the three methods are different with regard to error estimation. For the estimation, the choice of parameter space is important. Mulchrone et al. (2003) used the ϕ-R space (Fig.1) for bootstrap error estimation, where ϕ and R are the major-axis orientation and aspect ratio of an ellipse. This space has inconveniences. The paired data $\{R, \phi\}$ does not have one-to-one correspondence with points in the space: All points on the ϕ-axis represent circles. Points at the centers of the ellipses C and D in Fig.1 represent the same ellipses. In addition, distances between the ellipses A and B is the same with that between A′ and B′ in this diagram, but the dissimilarity of A and B is larger than that of A′ and B′. The significance of difference in ϕ become larger with increasing R. Is the dissimilarity between the ellipses A and B greater than that between A and C? The quantification of dissimilarity or distance is needed for the error estimation. An ordinary Rf/ϕ plot (Ramsay and Huber, 1983, p. 83), which uses $\log R$ instead of R itself for a coordinate, is inappropriate as well.

The parameter space of the method of Shimamoto and Ikeda (1976) is not convenient. We show that the choice of their parameter space results in the anisotropic resolution the method. The nature of anisotropy is demonstrated in this article.

As usual in Rf/ϕ strain analysis, volume changes associated with strain are ignored (Appendix A). Accordingly, we consider only the ellipses that have the same area with a unit circle, and deal with area-preserving strains.

Area-preserving strains are represented by matrices of determinant one (e.g., Matthews et al., 1974; Pollard and Fletcher, 2005). Ellipses with the area of π are represented also by those called shape matrices (e.g., Shimamoto and Ikeda, 1976; Wheeler, 1984). An identity matrix stands for a unit circle and null strain. Any strain ellipse has its reciprocal strain ellipse that is also represented a symmetric matrix with determinant one. Therefore, the action of strain and its objects (ellipses)

*Phone: +81 75 753 4266; fax: +81 75 753 4189.
Email address: yamaji@kueps.kyoto-u.ac.jp (Atsushi Yamaji)

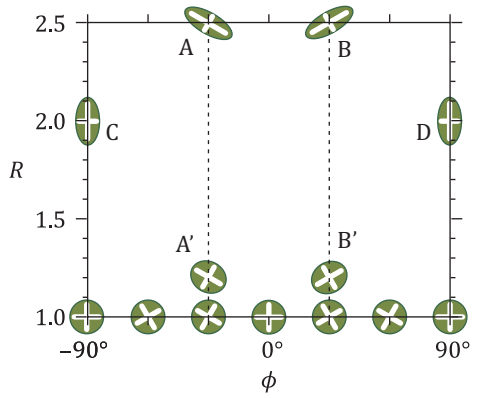

Figure 1: The Cartesian coordinates, ϕ and R for strain analysis, where ϕ and R are the major-axis orientation and aspect ratio of an ellipse. This diagram is inappropriate for error estimation in strain analysis, because dissimilarity between ellipses is not proportional to the distance between the points corresponding to the ellipses.

are all represented by matrices of determinant one, which make up a mathematical group termed special linear group. Accordingly, the theories of strain analysis can take advantage of group theory, thereby the theories have consistent geometrical interpretations. Specifically, special linear group is represented by the Lorentz transformation of points on a curved surface with a constant curvature of -1 in a Minkowski space (Ratcliffe, 2006). Some of the interpretations were given in Yamaji (2008), including the equivalence of the formulas of strain analysis (Dunnet, 1969; Lisle, 1985) with trigonometry on the surface. Additional ones are presented in this article. First, we introduce the space. Important symbols are listed in Table 1.

2. Method of Yamaji (2008)

2.1. Minkowski 3-space

As a parameter space for the shape and orientations of ellipses, we introduce the unit hyperboloid, H^2, denoted by $z = \sqrt{r^2+1}$, where r and z are cylindrical coordinates (Fig. 2). H^2 has rotational symmetry about the z-axis. By means of the rectangular Cartesian coordinates O-123, H^2 is denoted by

$$x_3^2 = x_1^2 + x_2^2 + 1 \quad (x_3 > 0), \tag{1}$$

where $\boldsymbol{x} = (x_1, x_2, x_3)^\top$ is a position vector. Using the parameter, ρ, this surface is expressed by the pair of equations, $r = \sinh\rho$ and $z = \cosh\rho$, because hyperbolic functions satisfy the identity, $\cosh^2\rho - \sinh^2\rho = 1$. Accordingly, the position vector of a point on H^2 has the expression,

$$\boldsymbol{x} = \begin{pmatrix} \sinh\rho\cos\psi \\ \sinh\rho\sin\psi \\ \cosh\rho \end{pmatrix}, \tag{2}$$

where ψ is the angle about the O-3 axis.

Now, we introduce the Minkowski norm,

$$\|\boldsymbol{x}\| = \sqrt{|\boldsymbol{x}\circ\boldsymbol{x}|}, \tag{3}$$

Table 1: List of symbols. The subscripts, 'i' and 'f,' indicate the quantities of pre- and post-strain ellipses, respectively; and 's' indicates the quantities of strain ellipse.

	Explanation	Ref.
$d_{\mathrm{H}}(\,)$	hyperbolic distance	Eq. (6)
$\boldsymbol{e}$	Position vectors with the endpoint on H^2	Eq. (8)
f	11-component of shape matrix	Eq. (11)
g	22-component of shape matrix	Eq. (11)
h	12- and 21-components of shape matrix	Eq. (11)
H^2	Unit hyperboloid	Eq. (1)
$\boldsymbol{N}$	Shape matrix	Eq. (11)
n	Number of ellipses	
R	Aspect ratio of an ellipse	
R_{s}	Aspect ratio of strain ellipse	
$\boldsymbol{s}$	Vector corresponding to shape matrix	Eq. (18)
$\boldsymbol{x}$	Position vectors with the endpoint on H^2	Eq. (2)
$\boldsymbol{\mu}$	Hyperbolic vector mean	Eq. (10)
ϕ	Major-axis orientation of an ellipse	
ϕ_{s}	Major-axis orientation of an ellipse	
ρ	Radial coordinate on H^2	Eq. (7)
ψ	Tangential coordinate	Eq. (7)
$\circ$	Lorentizan inner product	Eq. (4)
$\cdot$	Ordinary (Euclidean) inner product	
$\lvert\ \rvert$	Euclidean norm	
$\lVert\ \rVert$	Minkowski norm	Eq. (3)

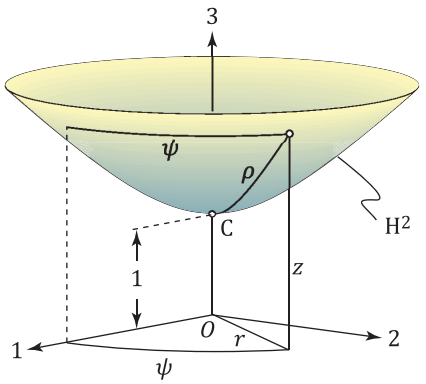

Figure 2: The unit hyperboloid, H^2, in a Minkowski 3-space, in which the rectangular Cartesian coordinates O-123 and cylindrical coordinates O-$r\psi z$ are defined. H^2 is denoted by $z = \sqrt{r^2+1}$. Each point on H^2 represents an ellipse. Point C at the base of this surface stands for a unit circle. C-$\rho\psi$ is the system of polar coordinates lying on H^2 just like latitude and longitude on the globe.

where

$$\boldsymbol{x}\circ\boldsymbol{y} = x_1y_1 + x_2y_2 - x_3y_3 \tag{4}$$

is known as the Lorentz inner product of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ (e.g., Ratcliffe, 2006). Since this product can have a negative value, the absolute value of $\boldsymbol{x}\circ\boldsymbol{x}$ is taken in Eq. (3). Due to the introduction of this norm, our parameter space becomes a Minkowski 3-space. H^2 is denoted as

$$\boldsymbol{x}\circ\boldsymbol{x} = -1, \tag{5}$$

which resembles the equation of a unit sphere $\boldsymbol{x}\cdot\boldsymbol{x} = 1$. As this analogy suggests, the formulas of spherical geometry have

their counterparts on H^2. Accordingly, spherical geometry is a useful guide to consider strain analysis in this space. The distance between the endpoints of the position vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is expressed as

$$d_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{y}) = \cosh^{-1}(-\boldsymbol{x} \circ \boldsymbol{y}), \tag{6}$$

called hyperbolic distance. Note the analogy of this equation with the angular distance on a unit sphere, $\cos^{-1}(\boldsymbol{x} \cdot \boldsymbol{y})$. It follows from Eq. (2) that $d_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{c}) = \cosh\rho$, where $\boldsymbol{c} = (0, 0, 1)^\top$ indicates the point at C in Fig. 2. Therefore, ρ is the hyperbolic distance from C. Accordingly, ρ and ψ are analogous to latitude and longitude on the globe. The parameters are regarded as the polar coordinates lying on H^2.

Given n elliptical strain markers, we have n corresponding points on H^2 referred to as $\boldsymbol{e}^{(1)}, \ldots, \boldsymbol{e}^{(n)}$. Their centroid is

2.2. Hyperbolic vector mean method

Let R and ϕ be the aspect ratio and major-axis orientation of an ellipse, respectively. We identify points on H^2 with ellipses through the equations,

$$\rho = \log R, \quad \psi = 2\phi, \tag{7}$$

where log means natural logarithm, ρ and ψ are the polar coordinates on H^2. It follows that an ellipse in the physical space is represented by the position vector,

$$\boldsymbol{e} = \begin{pmatrix} \sinh\rho\cos\psi \\ \sinh\rho\sin\psi \\ \cosh\rho \end{pmatrix} = \frac{1}{2}\begin{pmatrix} (R - 1/R)\cos 2\phi \\ (R - 1/R)\sin 2\phi \\ R + 1/R \end{pmatrix}. \tag{8}$$

There are one-to-one correspondences among $\{R, \phi\}$, $\{\rho, \psi\}$ and $\boldsymbol{e}$. A unit circle is indicated by the point C in Fig. 2, and long ellipses are represented by points far from C. If a unit circle becomes an ellipse with the aspect ratio R by strain, the principal radii of this ellipse are $\sqrt{R}$ and $1/\sqrt{R}$. Therefore, the corresponding logarithmic strain is $\log\sqrt{R} = \rho/2$. It means that ρ denotes doubled logarithmic strain.

$$\bar{\boldsymbol{e}} = \frac{1}{n}\left[\boldsymbol{e}^{(1)} + \cdots + \boldsymbol{e}^{(n)}\right]. \tag{9}$$

Then, the hyperbolic vector mean is defined as

$$\boldsymbol{\mu} = \bar{\boldsymbol{e}} \,/\, \|\bar{\boldsymbol{e}}\|. \tag{10}$$

Because of the convex downward shape of H^2 (Fig. 2), the centroid exists above H^2. The denominator in Eq. (10) drops a point at the centroid onto H^2. It means that $\boldsymbol{\mu} \circ \boldsymbol{\mu} = -1$. Yamaji (2008) showed that $\boldsymbol{\mu}$ represents the strain ellipse that best explains the n data. It is assumed for $\boldsymbol{\mu}$ to indicate the strain ellipse that the points on H^2 corresponding to the pre-strain ellipses in the physical space have the centroid at the point C (Section 6). The distribution of those points depicts the pre-strain shape fabric, which is discussed in Section 6.

Once this mean vector is obtained, the parameters of the strain ellipse is given by

$$R_{\mathrm{s}} = \exp\left(\sinh^{-1}\sqrt{\mu_1^2 + \mu_2^2}\right)$$

and

$$\phi_{\mathrm{s}} = \frac{1}{2}\tan^{-1}(\mu_2/\mu_1).$$

Let us consider the effect of the choice of reference orientation on the plane where strain markers are observed. Consider that two persons take different reference orientations, which meet at the angle of ϕ_0. Then, the points on H^2 indicated by $\boldsymbol{e}$-vectors (Eq. 8) of a person are rotated about the O-3 axis by the angle of $2\phi_0$ from those of the other person. Thanks to the rotational symmetry of H^2 about the O-3 axis, the relative positions of the points have no difference. It means that the strain ellipse estimated by the hyperbolic vector mean method does not depend on the choice of reference orientations.

3. Method of Shimamoto and Ikeda (1976)

3.1. Shimamoto-Ikeda method

Shimamoto and Ikeda (1976) identified the symmetric matrix,

$$\boldsymbol{N} = \begin{pmatrix} f & h \\ h & g \end{pmatrix}, \tag{11}$$

with an ellipse. $\boldsymbol{N}$ is called a shape matrix. The matrix should satisfy $\det\boldsymbol{N} = 1$ to indicate the ellipses that has the same area as the unit circle. Shimamoto and Ikeda (1976) showed that if $\boldsymbol{N}^{(1)}, \ldots, \boldsymbol{N}^{(n)}$ are the shape matrices of n elliptical strain markers, the matrix standing for the optimal strain ellipse is obtained by the mean,

$$\bar{\boldsymbol{N}} = \frac{1}{n}\left[\boldsymbol{N}^{(1)} + \cdots + \boldsymbol{N}^{(n)}\right], \tag{12}$$

where

$$\boldsymbol{N}^{(i)} = \begin{pmatrix} \cos\phi^{(i)} & -\sin\phi^{(i)} \\ \sin\phi^{(i)} & \cos\phi^{(i)} \end{pmatrix}\begin{pmatrix} 1/R^{(i)} & 0 \\ 0 & R^{(i)} \end{pmatrix}\begin{pmatrix} \cos\phi^{(i)} & \sin\phi^{(i)} \\ -\sin\phi^{(i)} & \cos\phi^{(i)} \end{pmatrix}. \tag{13}$$

The superscript (i) indicates the quantities of the ith ellipse. The explanation of Shimamoto and Ikeda (1976, p. 321) on the derivation of the principal radii of the strain ellipse is misleading, because they suggest as if the radii are obtained as the eigenvalues of $\bar{\boldsymbol{N}}^{-1}$. This is not true, because $\det\bar{\boldsymbol{N}} \neq 1$ (Wheeler, 1984). For this condition to be met, $\bar{\boldsymbol{N}}$ must be normalized as

$$\bar{\boldsymbol{N}}' = \bar{\boldsymbol{N}} / \sqrt{\det\bar{\boldsymbol{N}}}, \tag{14}$$

or equivalently,

$$\bar{\boldsymbol{N}}' = \frac{\boldsymbol{N}^{(1)} + \cdots + \boldsymbol{N}^{(n)}}{\sqrt{\det\left[\boldsymbol{N}^{(1)} + \cdots + \boldsymbol{N}^{(n)}\right]}}.$$

Then, we have $\det\bar{\boldsymbol{N}}' = 1$. The eigenvalues of $\bar{\boldsymbol{N}}'$ are equal to R_{s} and $1/R_{\mathrm{s}}$; the eigenvector corresponding to the larger eigenvalue indicates the major-axis of the optimal strain ellipse.

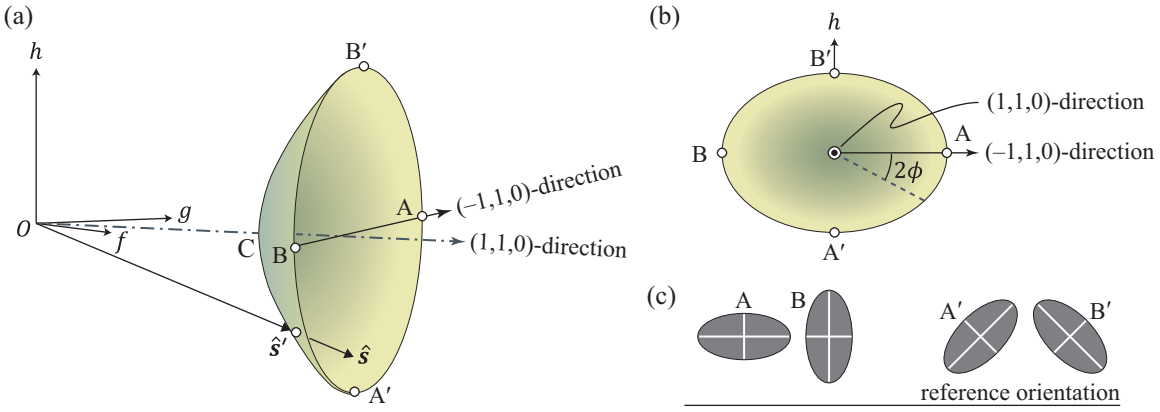

Figure 3: (a) The hyperboloid with elliptical aperture that is denoted by the equation $gh - h^2 = 1$, where O-fgh is rectangular Cartesian coordinates. Dot-dash line indicates the axis of the hyperboloid in the (1,1,0)-direction. Open circles depict points on the hyperboloid. (b) The ellipse lying on the plane perpendicular to this direction. (c) The elliptical markers, A, B, A′ and B′, have the aspect ratio of 2. Principal axes of the first two ellipses are parallel or perpendicular to the reference orientation, whereas those of remaining ones are inclined at 45°. The ellipses are represented by the points A, B, A′ and B′ in (a) and (b).

3.2. *Relationship of Shimamoto-Ikeda and the hyperbolic vector mean methods*

Wheeler (1984, Eq. A7) wrote the components of $\boldsymbol{N}$ in terms of R and ϕ:

$$f = \frac{1}{2}\left(\frac{1}{R} + R\right) + \frac{1}{2}\left(\frac{1}{R} - R\right)\cos 2\phi \tag{15}$$

$$g = \frac{1}{2}\left(\frac{1}{R} + R\right) - \frac{1}{2}\left(\frac{1}{R} - R\right)\cos 2\phi \tag{16}$$

$$h = \frac{1}{2}\left(\frac{1}{R} - R\right)\sin 2\phi. \tag{17}$$

We can think of these equations as the parametric expression of a curved surface in the fgh-space, where $R \geq 1$ and ϕ are the parameters (Fig. 3a). It is a hyperboloid with the axis in the $(1, 1, 0)$-direction. This surface is denoted by the equation, $\det \boldsymbol{N} = fg - h^2 = 1$. Point C at (1,1,0) stands for a unit circle. The aperture of the hyperboloid is elliptical with the aspect ratio of $\sqrt{2}$ (Fig. 3b). The major-axis is in the $(1, -1, 0)$ orientation. Increasing ϕ carries a point along this ellipse lying on the plane perpendicular to the axis: The ellipses, A, B, A′ and B′ in Fig. 3c are represented by the points A, B, A′ and B′ in Figs. 3a and b.

Now, we introduce the position vector,

$$\boldsymbol{s} = (f, g, h)^\top, \tag{18}$$

which has a one-to-one correspondence with $\boldsymbol{N}$ in Eq. (11). Then, the endpoint of $\boldsymbol{s}$ exists on the hyperboloid in Fig. 3. Let $\boldsymbol{s}^{(1)}, \ldots, \boldsymbol{s}^{(n)}$ be the vectors representing n elliptical strain markers. Then, their centroid,

$$\bar{\boldsymbol{s}} = \frac{1}{n}\left[\boldsymbol{s}^{(1)} + \cdots + \boldsymbol{s}^{(n)}\right],$$

corresponds to $\bar{\boldsymbol{N}}$, but the endpoint of this does not exist on the elliptical hyperboloid that is illustrated in Fig. 3. Instead, that of the vector

$$\bar{\boldsymbol{s}}' = \frac{\bar{\boldsymbol{s}}}{\bar{s}_1\bar{s}_2 - \bar{s}_3^2} \tag{19}$$

exists on the curved surface, where $\bar{s}_1$, $\bar{s}_2$ and $\bar{s}_3$ are the components of $\bar{\boldsymbol{s}}$. This vector represents the optimal strain ellipse. The denominator in the right-hand side of Eq. (19) equals the determinant of $\bar{\boldsymbol{N}}$. Due to the concave shape of the hyperboloid in the (1,1,0)-direction, the endpoint of $\bar{\boldsymbol{s}}$, which is surrounded by those of $\boldsymbol{s}^{(1)}, \ldots, \boldsymbol{s}^{(n)}$, does not exist on the hyperboloid, but in the right of the surface in Fig. 4. In other words, the denominator is greater than 1. It follows that $\det \bar{\boldsymbol{N}} > 1$. It means that the aspect ratio of the strain ellipse is overestimated if it is determined from $\bar{\boldsymbol{N}}$ insterad of $\bar{\boldsymbol{N}}'$ (Eq. 14).

Shimamoto and Ikeda's (1976) formulation is related with the Minkowski 3-space as follows. Using Eq. (8), Eqs. (15)–(17) are rewritten as

$$f = \cosh\rho - \sinh\rho\cos\psi \tag{20}$$

$$g = \cosh\rho + \sinh\rho\cos\psi \tag{21}$$

$$h = -\sinh\rho\sin\psi, \tag{22}$$

which are equivalent with Eq. (A9) of Wheeler (1984) except for the signs of the terms including sinh. This difference comes from the difference in the definitions of $\boldsymbol{N}^{(i)}$ in Eq. (13). It follows from Eqs. (2) and (20)–(22) that

$$x_1 = \frac{-f + g}{2} \tag{23}$$

$$x_2 = -h \tag{24}$$

$$x_3 = \frac{f + g}{2}. \tag{25}$$

Eqs. (23)–(25) say that O-fhg is a left-hand system with the coordinate axes inclined with respect to O-123 (Fig. 4). That is, the elliptical hyperboloid is mapped to H^2: Points on the surfaces have one-to-one correspondences through those equations. The factors of 1/2 in Eqs. (23) and (25) result in the elliptical aperture of Shimamoto and Ikeda's hyperboloid. Note the difference in the lengths of the arrows in Fig. 4: The arrows indicating $\boldsymbol{f}$ and $\boldsymbol{g}$ are longer than those of $\boldsymbol{h}$, 1, 2 and 3 by a factor of $\sqrt{2}$, resulting from the factors of 1/2.

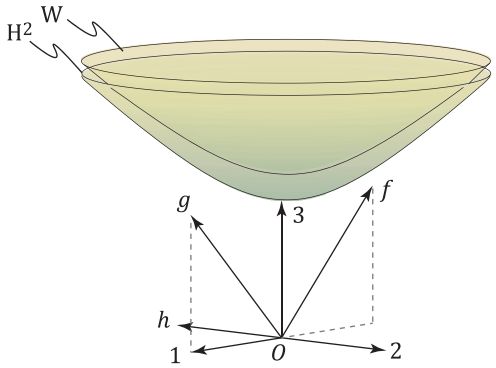

Figure 4: Hyperboloids, H^2 and W, the latter of which we call the Wheeler hyperboloid (Section 5). Components of Shimamoto and Ikeda's shape matrix define the coordinates O-fgh, the f- and g-axes of which are inclined and different scales from the Cartesian coordinate axes of the Minkowski 3-space.

In general, matrix components are affected by coordinate rotations. The functions of matrix components that are unaffected by the rotations are called invariants: They include

$$\text{trace}\,\boldsymbol{N} = N_{11} + N_{22} = f + g \tag{26}$$

$$\det \boldsymbol{N} = N_{11}N_{22} - N_{12}N_{21} = fg - h^2 \tag{27}$$

$$\text{Frob}\,\boldsymbol{N} = \sqrt{N_{11}^2 + N_{12}^2 + N_{21}^2 + N_{22}^2} = \sqrt{f^2 + g^2 + 2h^2}. \tag{28}$$

The last one is known as Frobenius norm (e.g., Meyer, 2000). Combining Eqs. (26) and (25), we have

$$\text{trace}\,\boldsymbol{N} = 2x_3. \tag{29}$$

Combining Eqs. (3), (4), (23)–(25) and (27), we have

$$\det \boldsymbol{N} = -x_1^2 - x_2^2 + x_3^2 = -\boldsymbol{x} \circ \boldsymbol{x} = \|\boldsymbol{x}\|^2. \tag{30}$$

It means that points on H^2 satisfying Eq. (5) stand for the matrices with determinant 1, indicating the correspondence of the elliptical hyperboloid in Fig. 3 and H^2. Frobenius norm (Eq. 28) becomes

$$\text{Frob}\,\boldsymbol{N} = \sqrt{2\left(x_1^2 + x_2^2 + x_3^2\right)} = \sqrt{2}|\boldsymbol{x}|,$$

where $|\boldsymbol{x}| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$ is the Euclidean norm of $\boldsymbol{x}$.

Because of the linearity of the transformation between the coordinate systems O-123 and O-fgh, the points denoted by $\overline{\boldsymbol{s}}$ and $\overline{\boldsymbol{e}}$ are mapped from each other via Eqs. (23)–(25). The denominator in Eqs. (10) is interchangeable with that in (19) via Eqs. (27) and (30). In other words, the strain ellipses determined by the method of Shimamoto and Ikeda (1976) and the hyperbolic vector mean method are mathematically identical.

4. Method of Mulchrone et al. (2003)

The mathematical equivalence of the method of Mulchrone et al. (2003) and the hyperbolic vector mean method was pointed out by Yamaji (2008). Accordingly, we explain the equivalence

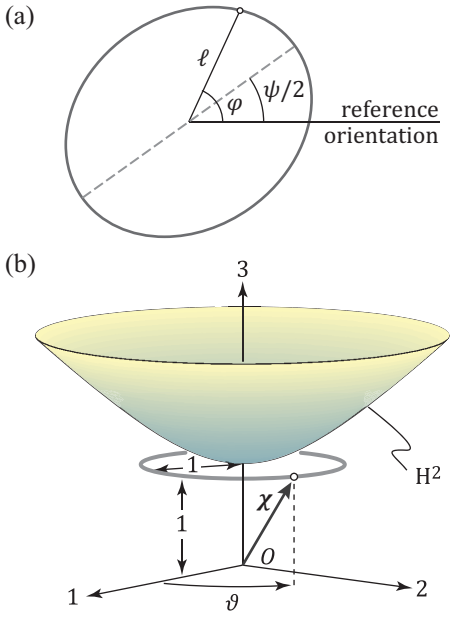

Figure 5: (a) Moving radius, $\ell(\varphi)$, of an ellipse. (b) Schematic picture for the explanation of the equivalence of the method of Mulchrone et al. (2003) and the hyperbolic vector mean method. Gray line depicts the unit circle, a point on which is indicated by $\boldsymbol{\chi} = (\cos\vartheta, \sin\vartheta, 1)^\top$.

briefly, here. Mulchrone et al. (2003) used the parametric expression of an ellipse,

$$\ell^{-2} = -\sinh\rho\cos\psi\cos 2\varphi - \sinh\rho\sin\psi\sin 2\varphi + \cosh\rho, \tag{31}$$

where $\ell(\varphi)$ is the moving radius of the ellipse (Fig. 5a); the ellipse is represented by ρ and ψ, the polar coordinates on H^2. They showed that the generalized mean with the power -2 of the n ellipses,

$$\ell_{\mathrm{s}}^{-2} \equiv \frac{1}{n}\left\{\left[\ell^{(1)}\right]^{-2} + \cdots + \left[\ell^{(n)}\right]^{-2}\right\}, \tag{32}$$

indicates the ellipse that is similar to the strain ellipse, where $\ell^{(i)}(\varphi)$ is the moving radius of the ith ellipse ($i = 1, 2, \ldots\ n$).

The relationship of this method and the hyperbolic vector mean method is shown as follows. In terms of position vectors in the Minkowski 3-space, Eq. (31) is rewritten as

$$\ell^{-2} = -\boldsymbol{e} \circ \boldsymbol{\chi}, \tag{33}$$

where $\boldsymbol{e} = (\sinh\rho\cosh\psi, \sinh\rho\sin\psi, \cosh\rho)^\top$ is the vector defined by Eq. (8), and $\boldsymbol{\chi} = (\cos\vartheta, \sin\vartheta, 1)^\top$ is the position vector of a point on the unit circle that is characterized by $z = 1$ and $r = 1$ (Fig. 5b). Using Eqs. (10) and (33), the right-hand side of Eq. (32) is rewritten as $-\left\|\boldsymbol{e}^{(1)} + \cdots + \boldsymbol{e}^{(n)}\right\| \boldsymbol{\mu} \circ \boldsymbol{\chi}$. Minkowski norm is a scalar quantity with a positive sign. Therefore, we have

$$\ell_{\mathrm{s}}^{-2} \propto -\boldsymbol{\mu} \circ \boldsymbol{\chi}. \tag{34}$$

It follows that the strain ellipses determined by this method are equal to those obtained by the hyperbolic vector mean method.

The strain ellipse determined by the method of Mulchrone et al. (2003) is not affected by the choice of reference orientations,

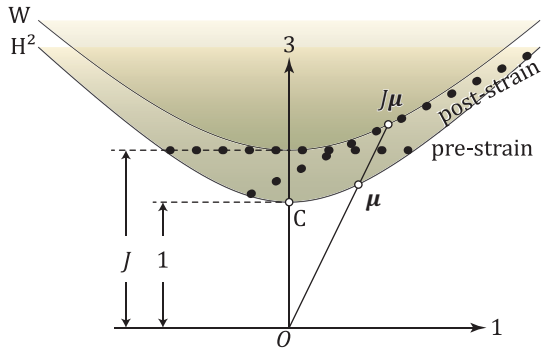

Figure 6: Hyperboloids, H^2 and the Wheeler hyperboloid, W. Closed circles indicate the points on H^2 corresponding to the pre- and post-strain ellipses.

because the freedom of this choice, ϕ_0, gives rise only to the rigid-body rotation of $\boldsymbol{\mu}$ and $\boldsymbol{\chi}$ about the O-3 axis by the angle of $2\phi_0$. The relative position indicated by these vectors are not affected by the rotation. What the formulation of this method is not endowed with is the measure of dissimilarity between ellipses, which enables error estimation.

5. Isometry of H^2

It is known that area-preserving strain is represented by the Lorentz transformation of points in the Minkowski 3-space (Ratcliffe, 2006). That is, the hyperbolic distances between points on H^2 corresponding to elliptical stain markers are preserved during strain, analogous to the preservation of great-circle distances during Euler rotations on the globe. Such a movement is termed isometry. Therefore, it is mathematically natural to use the Minkowski 3-space as the parameter space of strain analysis: Straining and destraining are represented by linear mapping and its inverse in the space (Yamaji, 2008). We have seen that the formulas of shape matrices have geometrical interpretations in the space, thereby theories of Rf/ϕ strain analysis were shown to be mathematically identical. Those are natural results of the introduction of the space. The hyperbolic vector mean method is based on the isometry. In addition, hyperbolic distance (Eq. 6) is a natural measure of distance or dissimilarity between ellipses, because the hyperbolic norms of the residuals, $\boldsymbol{\mu} - \boldsymbol{e}^{(i)}$ ($i = 1, 2, \ldots, n$), equal the doubled logarithmic strain needed to transform a unit circle to the ith ellipse. That is, the residuals have well-defined physical meanings concerning strain.

In the following part of this section, we introduce a theorem of Wheeler that allows us to evaluate dispersion in the shapes and attitudes of ellipses. Wheeler (1984) pointed out that the spread of pre-strain ellipses was preserved during strain. This theorem is obvious in the Minkowski 3-space, because area-preserving strain is represented by isometry of H^2. We explain that the theorem has a corresponding solid figure in the Minkowski 3-space, which we call the Wheeler hyperboloid (Fig. 6).

If $\bar{\boldsymbol{N}}_\mathrm{i}$ and $\bar{\boldsymbol{N}}_\mathrm{f}$ are the mean shape matrices (Eq. 12) for the pre- and post-strain ellipses, respectively, Wheeler (1984) derived

$$\det \bar{\boldsymbol{N}}_\mathrm{i} = \det \bar{\boldsymbol{N}}_\mathrm{f}, \tag{35}$$

which is called the distribution spread invariant (Wheeler, 1984). It follows from Eq. (30) that Eq. (35) is equivalent with

$$\|\bar{\boldsymbol{e}}_\mathrm{i}\| = \|\bar{\boldsymbol{e}}_\mathrm{f}\|, \tag{36}$$

Isotropic distribution of the shape and attitudes of pre-strain ellipses is represented by the points with the centroid, $\bar{\boldsymbol{e}}_\mathrm{i}$, at the point $(0, 0, J)$ on the O-3 axis (Fig. 6). Due to the concave upward shape of H^2, this point is above H^2, meaning that $J > 1$. This inequality was derived by Wheeler (1984). Eq. (36) says that post-strain centroid $\bar{\boldsymbol{e}}_\mathrm{f}$ exists always on the hyperboloid that has the base at the point $(0, 0, J)$. We call it the Wheeler hyperboloid. The denominator in Eq. (10) to yield the hyperbolic vector mean equals the right-hand side of Eq. (36) equals J (Fig. 6).

It is obvious that the isometry preserves the areas of polygons on H^2, the vertices of which correspond to ellipses in the physical space. Fig. 7 shows the equal-area projection of H^2 (Reynolds, 1993) to visualize this area preservation. The progressive strain of ellipses does not affect the area surrounded by the data points corresponding to the ellipses. Yamaji (2008) showed the orthographic and gnomonic projections have significance as well in strain and vorticity analyses. H^2 is not only the parameter space where the optimal strain and its error are evaluated, but also connected seamlessly with their visualization techniques.

6. Pre-strain fabric

The hyperbolic vector mean method assumes that the points on H^2 corresponding to the pre-strain elliptical markers have the centroid at the point C in Fig. 2. Coaxial deformation determined by the Rf/ϕ strain analysis is denoted by the linear transformation,

$$\boldsymbol{e}_\mathrm{f} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi_\mathrm{s} & -\sin\psi_\mathrm{s} \\ 0 & \sin\psi_\mathrm{s} & \cos\psi_\mathrm{s} \end{pmatrix} \begin{pmatrix} \cosh\rho_\mathrm{s} & \sinh\rho_\mathrm{s} & 0 \\ \sinh\rho_\mathrm{s} & \cosh\rho_\mathrm{s} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi_\mathrm{s} & \sin\psi_\mathrm{s} \\ 0 & -\sin\psi_\mathrm{s} & \cos\psi_\mathrm{s} \end{pmatrix} \boldsymbol{e}_\mathrm{i},$$

in the Minkowski 3-space, where ρ_s and ψ_s are the quantities of strain ellipse, $\boldsymbol{e}_\mathrm{i}$ and $\boldsymbol{e}_\mathrm{f}$ are the position vectors standing for the pre- and post-strain ellipses, respectively (Yamaji, 2008, Eq. 13). It is obvious from Fig. 7 that destraining is denoted by

$$\boldsymbol{e}_\mathrm{i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi_\mathrm{s}+\pi) & -\sin(\psi_\mathrm{s}+\pi) \\ 0 & \sin(\psi_\mathrm{s}+\pi) & \cos(\psi_\mathrm{s}+\pi) \end{pmatrix} \begin{pmatrix} \cosh\rho_\mathrm{s} & \sinh\rho_\mathrm{s} & 0 \\ \sinh\rho_\mathrm{s} & \cosh\rho_\mathrm{s} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi_\mathrm{s}+\pi) & \sin(\psi_\mathrm{s}+\pi) \\ 0 & -\sin(\psi_\mathrm{s}+\pi) & \cos(\psi_\mathrm{s}+\pi) \end{pmatrix} \boldsymbol{e}_\mathrm{f}. \tag{37}$$

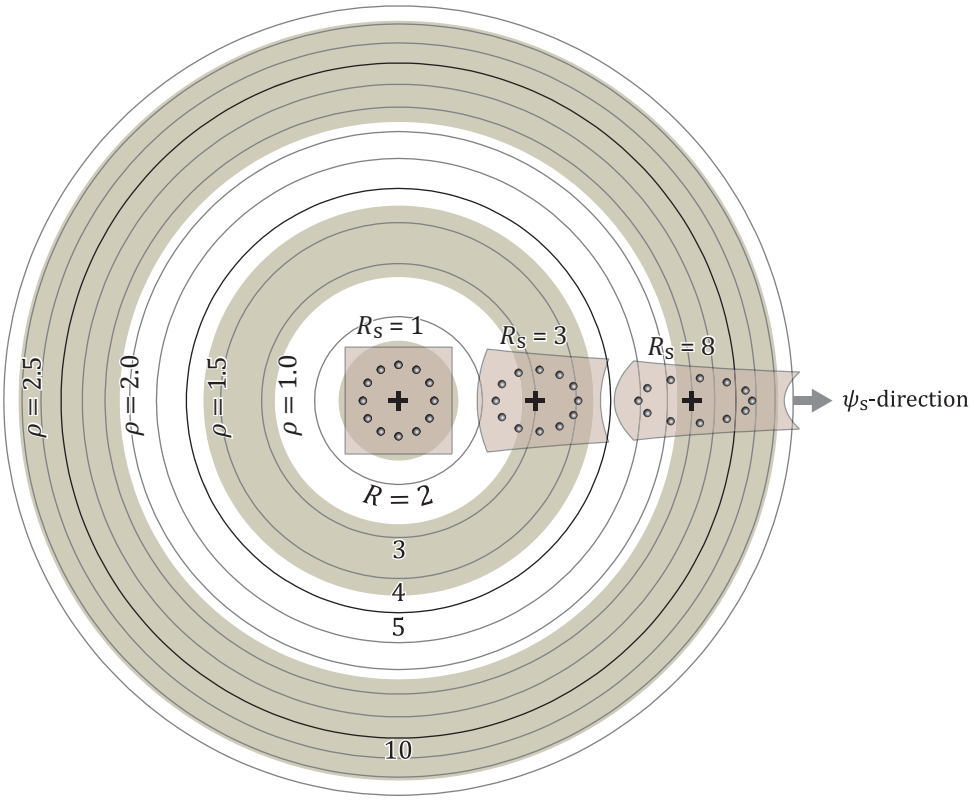

Figure 7: Equal-area projection of H^2 showing progressive strain of such ellipses that the corresponding points at the pre-strain stage make up a circle about the point C in Fig. 2. Crosses represent the strain ellipses with $R_s = 1$, 3 and 8. Arrow indicates the maximum stretching orientation. The area on H^2 surrounded by the points corresponding to ellipses is preserved by progressive strain. Concentric circles depicted by solid lines indicate iso-R lines, whereas concentric colored zones indicate the ranges of ρ. The distribution of data points on the equal-area projection changes through progressive strain on this plot from pre-strain circular distribution to oval ones. However, the variations are the result of the distortion of the equal-area projection, which is inevitable like the distortion of cartographic projection of the globe. The distribution does not change on H^2. Data points to the right of the cross for each cases of $R_s = 3$ and 8 are somewhat clustered than those to the left. This is also a result of the distortion.

Therefore, once strain ellipse is determined, the points corresponding to pre-strain ellipses are readily obtained through Eq. (37). Let $\boldsymbol{e}_i^{(1)}, \ldots, \boldsymbol{e}_i^{(n)}$ be the position vectors of those points. Then, we have their hyperbolic vector mean,

$$\boldsymbol{\mu}_i = \frac{\boldsymbol{e}_i^{(1)} + \cdots + \boldsymbol{e}_i^{(n)}}{\| \boldsymbol{e}_i^{(1)} + \cdots + \boldsymbol{e}_i^{(n)} \|}.$$

The hyperbolic vector mean method assumes $\boldsymbol{\mu}_i = (0, 0, 1)^\top$.

The equal-area projection of the points shows the shape fabric of the pre-strain ellipses. The hyperbolic vector mean method assumes that those points have the centroid at the point C in Fig. 2. That is, the method allows strain markers to have a variety of pre-strain shape fabric, but the deviation of the centroid of the pre-strain points from C gives rise to inaccuracy.

Not only the deviation, but also the distribution of the points denoted by $\boldsymbol{e}_i^{(1)}, \ldots, \boldsymbol{e}_i^{(n)}$ indicates the reliability of the result of the analysis. If pre-strain strain markers had isotropic fabric, their corresponding points on H^2 make a circular cluster around the point represented by $\boldsymbol{\mu}_i$ (Fig. 8a). The hyperbolic vector mean method requires only the coincidence of this point and C, i.e., $\boldsymbol{\mu}_i = (0, 0, 1)^\top$. This condition is met not only by the circular cluster but also by various types of clusters indicating anisotropic fabrics of pre-strain ellipses. Points in the elliptical cluster in Fig. 8b have the corresponding ellipses that have dominant major-axis orientations at $\phi = \pm 45°$. Such bimodal distributions are often found in depositional fabric, which suggests paleocurrent directions (Potter and Pettijohn, 1963, p. 44). The points in Fig. 8c have the centroid also at C, but make three clusters. Such an odd pattern suggests inaccuracy of the strain determined by the Rf/ϕ strain analysis: For example, the strain markers of different clusters may be the formed in different deformation stages. In this respect, the equal-area projection of pre-strain points, $\boldsymbol{e}_i^{(1)}, \ldots, \boldsymbol{e}_i^{(n)}$, is important to check the validity of the Rf/ϕ strain analysis.

Shimamoto and Ikeda (1976) and Mulchrone et al. (2003) wrote that their algebraic method worked for the case of the circular cluster in Fig. 8a. However, the three algebraic methods are mathematically identical. Therefore, the above argument applies not only to the hyperbolic vector metan method but also to their method.

7. Error estimation

We have seen the equivalence of results of the methods of Shimamoto and Ikeda (1976), Mulchrone et al. (2003) and Ya-

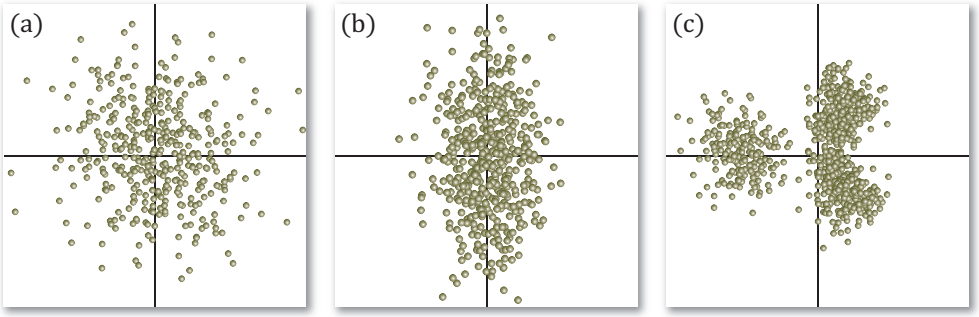

Figure 8: Schematic illustrations for the distribution of points on H^2 with the centroid at the point C, which is depicted by a cross. (a) Distribution with rotational symmetry about C, indicating isotropic shape fabric of the corresponding ellipses in the physical space. (b) An elliptical distribution.(c) More complicated distribution indicating anisotropic shape fabric of the corresponding ellipses.

maji (2008). However, they show difference in the difficulty of error estimation.

For this estimation, distance or dissimilarity between ellipses with different shapes and attitudes have to be defined quantitatively, because the confidence limit of the optimal strain depends on the spread in the shapes and attitudes of strain markers (e.g., Yamaji, 2005). The spread is evaluated from the distances of data points from the point corresponding to the optimal strain ellipse. The distance measure and the parameter space where the measure is defined are appropriate for this purpose, if they have satisfy the conditions:

1. The paired data of a marker $\{R, \phi\}$ have to have a one-to-one correspondence with a point in the parameter space.
2. The distances between the points representing elliptical strain markers must not be affected the choice of reference orientation on the plane where the markers are observed.

H^2 as the parameter space for ellipses meets these demands. Bootstrap error estimation was made on H^2 by Yamaji (2008).

The formulation of Shimamoto and Ikeda (1976) is not convenient for error estimation, because it does not satisfy the second condition. Shape matrices are denoted by points in the fgh-space. Four ellipses, A, A′, B and B′, in Fig. 3c have the same aspect ratios, but have different attitudes. The ellipses A and B have difference only in the ϕ values by 90°. The same is true for the ellipses A′ and B′. However, the Euclidean distances $\overline{\mathrm{AB}}$ and $\overline{\mathrm{A'B'}}$ are not equal to each other (Fig. 3b).

The method of Shimamoto and Ikeda (1976) is not suitable for error estimation owing to its anisotropic resolution: The dispersion of ellipses evaluated in the fgh-space depends on the choice of reference orientation. To show this anisotropy, we conducted a simple numerical experiment. Nine ellipses with tiny difference in R and ϕ values were assumed (Fig. 9a). Specifically, assumed apect ratios and major-axis orientations were 1.1 ± 0.001 and $\phi_0 \pm 0.01°$; and the dispersion of those ellipses were observed during the increase in ϕ_0. The dispersion was evaluated as follows. First, the nine points corresponding to the ellipses in the fgh-space were calculated. Let $\boldsymbol{s}^{(1)}, \ldots, \boldsymbol{s}^{(9)}$ be the position vectors indicating the points (Eq. 18). Then,

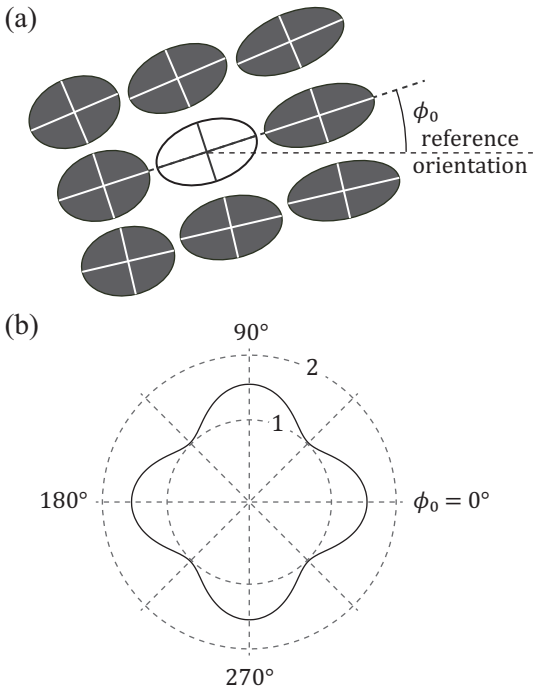

Figure 9: Anisotropic resolution of the Shimamoto-Ikeda method. (a) Schematic illustrations of ellipses with slightly different aspect ratios and major-axis orientations. The ellipse at the middle has the aspect ratio of 1.1, and major-axis orientation, ϕ_0. The ratios of the ellipses are exaggerated. (b) Polar plot showing the dispersion, σ, of the nine ellipses in (a) in the fgh-space. Curves show σ versus ϕ_0, where σ is normalized by the size of error ellipsoid of $\bar{\boldsymbol{e}}$ in the Minkowski 3-space. See text in detail.

the matrix,

$$\boldsymbol{C} = \frac{1}{9} \sum_{i=1}^{9} \left[\boldsymbol{s}^{(i)} - \bar{\boldsymbol{s}}\right] \left[\boldsymbol{s}^{(i)} - \bar{\boldsymbol{s}}\right]^{\top}$$

is the covariance matrix (e.g., Johnson and Wichern, 2002) of

the points. The dispersion σ was evaluated by

$$\sigma = \sqrt{\sum_{i,j=1}^{3} C_{ij}^2},$$

the right-hand side of which is the Frobenius norm of $\boldsymbol{C}$. As a result, σ represents the size of error ellipsoid of $\bar{\boldsymbol{s}}$ in the fgh-space. The norm can be thought of as the magnitude of $\boldsymbol{C}$. Fig. 9b shows the dispersion versus ϕ_0. Since the dispersion evaluated in the same manner from the $\boldsymbol{e}$ vectors defined by Eq. (8) does not have such ϕ_0-dependence. Therefore, the dispersion of the $\boldsymbol{s}$ vectors was normalized by that of the $\boldsymbol{e}$ vectors to show the anisotropy. Solid circle in Fig. 9b shows the latter dispersion.

As a result, the dispersion showed oscillatory behavior with the periodicity of 90° (Fig. 9b). The ratio of maximum and minimum dispersion was $\sqrt{2}$. Accordingly, the method of Shimamoto and Ikeda (1976) overestimates the dispersion of ellipses if their major-axes are subparallel to the reference orientation, whereas underestimation occurs if the major-axes meet a reference orientation at around 45°. In other words, the method has the best resolution when the reference orientation is parallel to a strain axis.

The ellipticity of the aperture of the hyperboloid gives rise to the anisotropy. Open circles in Fig. 10 depict the points representing the ellipses with the same aspect ratios and with the small intervals in the $\Delta\phi$ values. The central angles of the points have the differences of $2\Delta\phi$: Angles in the physical space are doubled in the fgh-space. Obviously, Euclidean distances between the points have the maximum values at $2\phi = 0$ and 180°, and the minimum values at $2\phi = 90$ and 270°, resulting in the ϕ-dependency of the dispersion.

In contrast, the methods of Mulchrone et al. (2003) and Yamaji (2008) do not have such dependency, so that the choice of reference orientation does not affect their resolution.

8. Conclusions

The algebraic methods of Rf/ϕ strain analysis by Shimamoto and Ikeda (1976), Mulchrone et al. (2003) and Yamaji (2008) were shown to yield the same strain ellipses. However, error estimation is difficult for the former two methods in their formulations. The Minkowski 3-space and the unit hyperboloid, H^2, are natural parameter space for strain analysis. Theorems of stain analysis have geometrical interpretations in the space, which also provides visualization techniques for Rf/ϕ data, the optimal strain and its confidence region.

9. Acknowledgments

We are grateful to R.J. Lisle and F.W. Vollmer for kind reviews. This works was supported financially by JSPS (22340151).

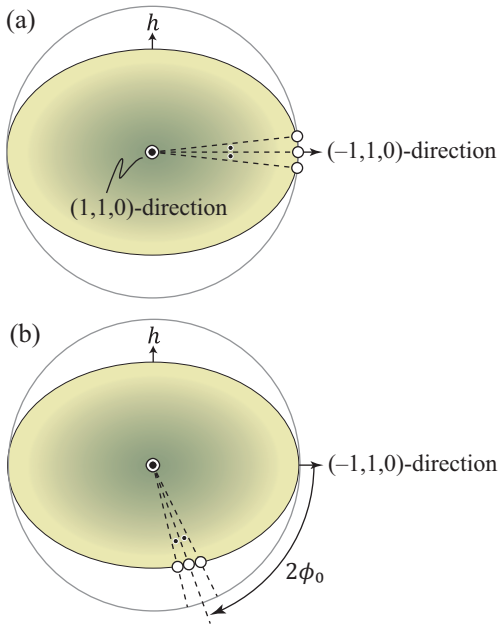

Figure 10: Illustrations for the origin of the anisotropic resolution of the Shimamoto-Ikeda method. Ellipses show the sections of the elliptical hyperboloid in the O-fgh parameter space (Fig. 3a). Open circles represent the three ellipses that has the same R values but has different ϕ values with constant angular intervals, which are indicated by solid circles. The situations of (a) and (b) are different only in their reference orientations. Note that the distances between the three points on the elliptical hyperboloid in (a) is greater than that in (b). This difference indicates the anisotropic resolution. (c) Polar plot showing the anisotropy. See text in detail.

Appendix A. Validity of ignoring volume changes

The methods of Rf/ϕ strain analysis, including the algebraic ones, are shown in this section to be able to deal with such strain markers that experienced volume changes. Let $\boldsymbol{D}$ be the deformation gradient tensor (e.g., Chadwick, 1999) for isochoric (volume-preserving) shape change, and let $\varsigma\boldsymbol{I}$ be the deformation gradient tensor for dilatation with no shape change, where $\boldsymbol{I}$ is the identity matrix, ς the ratio of expansion. Then, it can be seen that the order of the deformations is exchangeable to describe an arbitrary deformation, because $\boldsymbol{D}$ and $\boldsymbol{I}$ are commutable: $\boldsymbol{D}\varsigma\boldsymbol{I} = \varsigma\boldsymbol{I}\boldsymbol{D} = \varsigma\boldsymbol{D}$. Suppose that incremental deformation is decomposed into stepwise isochoric and dilating infinitesimal deformations, $\varsigma^{(1)}\boldsymbol{D}^{(1)}, \ldots, \varsigma^{(m)}\boldsymbol{D}^{(m)}$, where m is the number of steps (Fig. A1). Then, the total, finite deformation is denoted by the deformation gradient tensor of the form,

$$\left[\varsigma^{(m)}\boldsymbol{D}^{(m)}\right]\cdots\left[\varsigma^{(1)}\boldsymbol{D}^{(1)}\right] = \left[\varsigma^{(1)}\cdots\varsigma^{(m)}\boldsymbol{I}\right]\left[\boldsymbol{D}^{(m)}\cdots\boldsymbol{D}^{(1)}\right].$$

The left-hand side of this equation denotes the iteration of the pair of isochoric and dilating deformations, and the first and second brackets in the right-hand side are the volume and shape changes, respectively. This equation says that the series of isochoric deformations followed by the series of dilating ones has the same effect with the series of iterative ones. The series of dilations does not affect R_s and ϕ_s, i.e., the aspect ratio and major-axis orientation of the strain ellipse, of the strain denoted

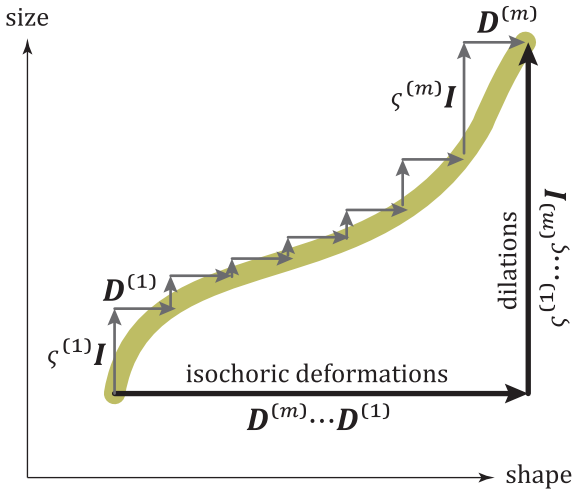

Figure .11: Schematic diagram for a deformation path (thick line) as a result of infinitesimal deformations (thin arrows). The horizontal axis denotes the shapes and attitudes of ellipses, and the vertical one denotes the size of ellipses. Thick solid arrows depict the two-step deformation that is equivalent with the total of the incremental deformations.

by $\boldsymbol{D}^{(n)}\cdots\boldsymbol{D}^{(1)}$. Therefore, we can estimate the strain ellipse for non-isochoric deformations by means of geological strain analysis methods.

The Rf/ϕ strain analysis deal only with the paired data, $\{R, \phi\}$, where the information of ellipse size is abstracted away. The analysis ignores the sizes of strain markers to determine R_s and ϕ_s at the cost of the indeterminacy of volume changes. Geological data other than the paired data of post-strain elliptical strain markers are necessary to determine volume changes.

References

Chadwick, P., 1999. Continuum Mechanics: Concise Theory and Problems, Second Edition. Dover Publications, Mineola.

Dunnet, D., 1969. A technique of finite strain analysis using elliptical particles. Tectonophysics 7, 117–136.

Dunnet, D., Siddans, A.W.B., 1971. Non-random sedimentary fabrics and their modification by strain. Tectonophysics 12, 307–325.

Johnson, R.A., Wichern, D.W., 2002. Applied Multivariate Statistical Analysis, 5th Edition. Prentice-Hall, Upper Saddle River.

Lisle, R.J., 1985. Geological Strain Analysis: A Manual for the Rf/ϕ Technique. Pergamon. Oxford.

Matthews, P.E., Bond, R.A.B., Van Den Berg, J.J., 1974. An algebraic method of strain analysis using elliptical markers. Tectonophysics 24, 31–67.

Meyer, C., 2000. Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia.

Mulchrone, K.F., O'Sullivan, F., Meer, P.A., 2003. Finite strain estimation using the mean radial length of elliptical objects with bootstrap confidence intervals. Journal of Structural Geology 25, 529–539.

Pollard, D.D., Fletcher, R.C., 2005. Fundamentals of Structural Geology. Cambridge University Press, Cambridge.

Potter, P.E., Pettijohn, F.J., 1963. Paleocurents and Basin Analysis. Springer, Berlin.

Ramsay, J.G., 1967. Folding and Fracturing of Rocks. McGraw-Hill, New York.

Ramsay, J.G., Huber M.I., 1983. The Techniques of Modern Structural Geology, Volume 1: Strain Analysis. Academic Press, London.

Ratcliffe, J.G., 2006. Foundations of Hyperbolic Manifolds, Second Edition. Springer Science and Business Media, New York.

Reynolds, W.F., 1993. Hyperbolic geometry on a hyperboloid. American Mathematical Monthly 100, 44–455.

Shimaoto, T., Ikeda, Y., 1976. A simple algebraic method for strain estimation from deformed eillipsoidal objects: 1. Basic theory. Tectonophysics 36, 315–337.

van Eeden, C., 2006. Restricted Parameter Space Estimation Problems: Admissibility and Minimaxity Properties. Springer, New York.

Wheeler, J., 1984. A new plot to display the strain of elliptical markers. Journal of Structural Geology 6, 417–423.

Yamaji, A., 2005. Finite tectonic strain and its error, as estimated from elliptical objects with a class of initial preferred orientations. Journal of Structural Geology 27, 2030–2042.

Yamaji, A., 2008. Theories of strain analysis from shape fabrics: A perspective using hyperbolic geometry. Journal of Structural Geology 30, 1451–1465.